Journal of Multivariate Analysis 102 (2011) 937–947

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Efficient empirical-likelihood-based inferences for the single-index model

Zhensheng Huang [a,*], Riquan Zhang [b]

[a] School of Mathematics, Hefei University of Technology, Hefei, Anhui, 230009, PR China

[b] Department of Statistics, East China Normal University, Shanghai, 200241, PR China

ARTICLE INFO

Article history:
Received 20 April 2010
Available online 22 February 2011

AMS 2000 subject classifications:
62G08
62E20

Keywords:
Confidence interval
Link function
Profile empirical likelihood
Single-index model
Single parameter

ABSTRACT

This article proposes the efficient empirical-likelihood-based inferences for the single component of the parameter and the link function in the single-index model. Unlike the existing empirical likelihood procedures for the single-index model, the proposed profile empirical likelihood for the parameter is constructed by using some components of the maximum empirical likelihood estimator (MELE) based on a semiparametric efficient score. The empirical-likelihood-based inference for the link function is also considered. The resulting statistics are proved to follow a standard chi-squared limiting distribution. Simulation studies are undertaken to assess the finite sample performance of the proposed confidence intervals. An application to real data set is illustrated.

1. Introduction

The single-index model has the following form:

$$Y = \eta(\beta^T X) + \varepsilon, \tag{1}$$

where $\eta(\cdot)$ is an unknown link function, $\beta \in R^p$ is an unknown parametric vector, for the sake of identifiability, we assume that $\|\beta\| = 1$ and the first component of β is positive, where $\|\cdot\|$ denotes the Euclidean norm, $Y \in R$, $X \in R^p$, and ε is the random error such that $E(\varepsilon|X) = 0$ almost surely. The appeal of the single-index model is that by focusing on an index $\beta^T X$, the so-called "curse of dimensionality" in fitting multivariate nonparametric regression functions is avoided. Because of its importance, much effort has been devoted to studying its estimation and other relevant inference problems. For example, see [8,7,1,15,19,16,12,9,21,17,14,13,2] and so on.

This paper mainly focuses on a relevant topic of the construction of confidence intervals of the parameter β and the link function $\eta(\cdot)$. A motivation of this study comes from an analysis of environmental data, consisting of daily measurements of pollutants and other environmental factors in Hong Kong between January 1, 2000 and December 31, 2000. The four variables—sulfur dioxide (in g/m^3) X_1, nitrogen oxides (in g/m^3) X_2, respirable suspended particulate (in g/m^3) X_3 and ozone (in g/m^3) X_4 are considered. Our main interests include two aspects: one is to examine the relationship between the levels of chemical pollutants and the number of daily total hospital admissions (Y) for respiratory diseases in Hong Kong. To avoid the curse of dimensionality, such a problem can be tackled by using the single-index model (1). Namely an index based on a linear combination of the pollutants $\beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$, which is called air pollution index (API), is

* Corresponding author.
E-mail addresses: stahzs@126.com, my2007email@sina.com (Z. Huang).

0047-259X/$ – see front matter
doi:10.1016/j.jmva.2011.01.011

estimated and used. The other is to construct the confidence intervals of every parameter β_l and the link function $\eta(\cdot)$. Here $l = 1, 2, 3, 4$. To this end, we will use the popular empirical likelihood method [10] to estimate the parameter β and the link function $\eta(\cdot)$. Furthermore, based on the estimates, the confidence intervals of the single parameter and the link function can be constructed. The detailed analysis of this data set will be reported in Section 3.

Although some authors has used the empirical likelihood method to construct the confidence regions of the parameter β in the existing literature (see [21,17,4]), their proposed method could not directly be employed to tackle the above proposed problems. The reasons are as follows: Firstly, even if they also used the usual 'delete-one-component' method to increase the accuracy of the proposed confidence region of β, this method is not efficient for the high-dimensional parameter β because this method reduced the dimension of β to $p - 1$, which is only one dimension less. For example, when $p \geqslant 3$ we cannot obtain the confidence interval of the single parameter of β, which is of primary interest in the analysis of the above environmental data. To improve the accuracy of empirical-likelihood-based region, we need to propose an efficient method to construct the confidence interval of the single component of β. Secondly, the empirical likelihood method proposed by Zhu and Xue [21] could not be used to construct the confidence interval of the link function $\eta(\cdot)$. In addition, to the best of our knowledge, we do not see specific works for the statistical inference about the confidence interval of the link function $\eta(\cdot)$ of the model (1). Thirdly, Zhu and Xue [21] did not considered the testing problem on the single parameter β_l, $l = 1, 2, 3, 4$, by using the empirical likelihood method. How to use the empirical likelihood method to select the significant variables of the API is also of interest.

In this paper, we propose an efficient profile empirical likelihood method to construct the confidence interval of the single component of β. To this purpose, we also propose an efficient maximum empirical likelihood estimator (MELE) $\hat{\beta}$ of the parameter β based on a semiparametric efficient score, which is motivated by the idea of Zhu and Xue [21]. By replacing some components of β with their estimators, an estimated profile empirical likelihood ratio statistic is constructed and it can be shown to be asymptotically chi-square distributed. Then the confidence interval of the single component of β is obtained. Furthermore, an empirical-likelihood-based test on the single component of β could be obtained by using the duality between confidence intervals and hypothesis tests. We also use the empirical likelihood method to construct the confidence interval of the link function $\eta(\cdot)$ in the model (1). Although a related work for the varying-coefficient model was given by Xue and Zhu [18], their used tools cannot directly be applied to the model (1) because of the different estimation methods in two models. Thus, our method is by no means straightforward. This motivates us to propose a new empirical likelihood method to construct the confidence interval of $\eta(\cdot)$.

The rest of this paper is organized as follows. In Section 2 an efficient profile empirical log-likelihood ratio for the single parameter of β is defined and a corrected empirical likelihood confidence interval for the link function $\eta(\cdot)$ is also constructed. Section 3 provides examples based on simulated and real data, and a comparison between the proposed empirical likelihood method and the normal approximation method is performed in term of coverage probabilities and widths of confidence intervals. The proofs of the main results are collected in Appendix.

2. Methodology and main results

2.1. Profile empirical likelihood for β

Suppose that $\{X_i, Y_i\}_{i=1}^n$ is a sample of size n from the model (1). Thus

$$Y_i = \eta(\beta^T X_i) + \varepsilon_i, \quad i = 1, \ldots, n. \tag{2}$$

Because $\|\beta\| = 1$ means that the true value of β is the boundary point on the unit sphere, $\eta(\beta^T X_i)$ does not have the derivative at the point β. For this, we use the "delete-one-component" method. The details are as follows. We assume that the true parameter β has a positive component (otherwise, consider $-\beta$). Without loss of generality, we assume $\beta_r > 0$, where β_r is the rth component of β for $1 \leqslant r \leqslant p$. For $\beta = (\beta_1, \ldots, \beta_p)^T$, let $\beta^{(r)} = (\beta_1, \ldots, \beta_{r-1}, \beta_{r+1}, \ldots, \beta_p)^T$ be a $p - 1$ dimensional parameter vector after removing the rth component β_r in β. Then we may write

$$\beta = \beta(\beta^{(r)}) = (\beta_1, \ldots, \beta_{r-1}, (1 - \|\beta^{(r)}\|^2)^{1/2}, \beta_{r+1}, \ldots, \beta_p)^T. \tag{3}$$

The value of β can be determined by $\beta^{(r)}$. Therefore, we need only to consider the confidence region of $\beta^{(r)}$. The true parameter $\beta^{(r)}$ must satisfy the constraint $\|\beta^{(r)}\| < 1$. Thus, β is infinitely differential in a neighborhood of $\beta^{(r)}$, and the Jacobian matrix is

$$\mathbf{J}_{\beta^{(r)}} = \frac{\partial \beta}{\partial \beta^{(r)}} = (\gamma_1, \ldots, \gamma_p)^T,$$

where $\gamma_s (1 \leqslant s \leqslant p, s \neq r)$ is a $p - 1$ dimensional vector with sth component 1, and $\gamma_r = -(1 - \|\beta^{(r)}\|^2)^{-1/2} \beta^{(r)}$.

Now we introduce an estimated auxiliary random vector for $\beta^{(r)}$ (see [21]

$$\hat{\xi}_i(\beta^{(r)}) = \{Y_i - \hat{\eta}(\beta^T X_i; \beta)\} \hat{\eta}'(\beta^T X_i; \beta) \mathbf{J}^T_{\beta^{(r)}} (X_i - \hat{\mu}(\beta^T X_i; \beta)), \tag{4}$$

where $\hat{\eta}(\beta^T X_i;\beta)$, $\hat{\eta}'(\beta^T \mathbf{X}_i;\beta)$ and $\hat{\mu}(\beta^T X_i;\beta)$ are the estimators of $\eta(\beta^T X_i)$, $\eta'(\beta^T X_i)$ and $\mu(\beta^T X_i)$, respectively, where $\mu(x) = E(X|\beta^T X = x)$. By the local linear method (see [5]), they can be defined as

$$\hat{\eta}(x;\beta) = \sum_{i=1}^{n} M_{ni}(x;\beta) Y_i, \qquad \hat{\eta}'(x;\beta) = \sum_{i=1}^{n} \widetilde{M}_{ni}(x;\beta) Y_i, \qquad \hat{\mu}(x;\beta) = \sum_{i=1}^{n} M_{ni}(x;\beta) X_i, \tag{5}$$

where $M_{ni}(x;\beta) = V_{ni}(x;\beta)/\sum_{j=1}^{n} V_{nj}(x;\beta)$, $\widetilde{M}_{ni}(x;\beta) = \widetilde{V}_{ni}(x;\beta)/\sum_{j=1}^{n} V_{nj}(x;\beta)$, and

$$V_{ni}(x;\beta) = K_h(\beta^T X_i - x)[S_{n,2}(x;\beta) - (\beta^T X_i - x) S_{n,1}(x;\beta)],$$
$$\widetilde{V}_{ni}(x;\beta) = K_h(\beta^T X_i - x)[(\beta^T X_i - x) S_{n,0}(x;\beta) - S_{n,1}(x;\beta)],$$
$$S_{n,k}(x;\beta) = \frac{1}{n}\sum_{i=1}^{n} (\beta^T X_i - x)^k K_h(\beta^T X_i - x), \quad k = 0, 1, 2,$$

where $K_h(\cdot) = K(\cdot/h)/h$, $K(\cdot)$ is a kernel function and h is a bandwidth.

Thus we can construct an estimated empirical log-likelihood ratio defined as

$$\widehat{\mathcal{R}}_n(\beta^{(r)}) = -2 \max \left\{ \sum_{i=1}^{n} \log(np_i) : p_i \geqslant 0, \sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} p_i \hat{\xi}_i(\beta^{(r)}) = 0 \right\}. \tag{6}$$

Zhu and Xue [21] have showed that the proposed $\widehat{\mathcal{R}}_n(\beta^{(r)})$ is asymptotically χ^2 distributed and by this the confidence region for $\beta^{(r)}$ can be constructed. But when the dimension $p \geqslant 3$ their method cannot be used to construct the confidence interval of the single component of the parameter β. To improve the accuracy of confidence region of β, we propose an efficient profile empirical likelihood method to construct a confidence interval for a single parameter of the parameter β. To this purpose, we need to obtain the efficient estimator $\hat{\beta}^{(r)}$ of $\beta^{(r)}$, which can be defined by maximizing $-\widehat{\mathcal{R}}_n(\beta^{(r)})$, and then obtain an estimator of β via a transform. This is called the maximum empirical likelihood estimator (MELE). From (2.9) and (A.1) of [21], we can easily see that the estimator $\hat{\beta}^{(r)}$ is, asymptotically, the solution to the efficient estimating equation $\sum_{i=1}^{n} \hat{\xi}_i(\beta^{(r)}) = 0$. The following result shows that $\hat{\beta}$ and $\hat{\beta}^{(r)}$ are asymptotically normal.

Theorem 1. *Suppose that Conditions* C_1–C_5 *in the Appendix hold. If* $\beta^{(r)}$ *is the true value of the parameter, then*

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \Sigma), \tag{7}$$

where $\xrightarrow{d}$ *stands for convergence in distribution,* $\Sigma = \mathbf{J}_{\beta^{(r)}} V^{-1} Q V^{-1} \mathbf{J}_{\beta^{(r)}}^T$, $Q = E\{\varepsilon^2 \eta'(\beta^T X)^2 \mathbf{J}_{\beta^{(r)}}^T (X - E(X|\beta^T X))(X - E(X|\beta^T X))^T \mathbf{J}_{\beta^{(r)}}\}$ *and* V *is defined in Condition* C_4.

Now by replacing β with $\hat{\beta}_{/s}$ in (4) we can define an estimated auxiliary random vector as

$$\hat{\xi}_{i,s}(\beta_s) = e_s^T \{Y_i - \hat{\eta}(\hat{\beta}_{/s}^T X_i;\hat{\beta}_{/s})\} \hat{\eta}'(\hat{\beta}_{/s}^T X_i;\hat{\beta}_{/s}) \mathbf{J}_{\hat{\beta}^{(r)}}^T (X_i - \hat{\mu}(\hat{\beta}_{/s}^T X_i;\hat{\beta}_{/s})), \tag{8}$$

where $\hat{\beta}_{/s} = \beta(\hat{\beta}_{/s}^{(r)})$, $\hat{\beta}_{/s}^{(r)} = (\hat{\beta}_1, \ldots, \beta_s, \hat{\beta}_{r-1}, \hat{\beta}_{r+1}, \ldots, \hat{\beta}_p)^T$, $s \in \{1, \ldots, r-1, r+1, \ldots, p\}$ and e_s is a $p-1$ dimensional vector with sth component 1. Then we can construct an estimated profile empirical log-likelihood ratio defined as

$$\widehat{\mathcal{R}}_{n,s}(\beta_s) = -2 \max \left\{ \sum_{i=1}^{n} \log(np_i) : p_i \geqslant 0, \sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} p_i \hat{\xi}_{i,s}(\beta_s) = 0 \right\}. \tag{9}$$

By the Lagrange multiplier method, $\widehat{\mathcal{R}}_{n,s}(\beta_s)$ can be represented as

$$\widehat{\mathcal{R}}_{n,s}(\beta_s) = 2 \sum_{i=1}^{n} \log\{1 + \lambda \hat{\xi}_{i,s}(\beta_s)\}, \tag{10}$$

where λ is determined by

$$\frac{1}{n} \sum_{i=1}^{n} \frac{\hat{\xi}_{i,s}(\beta_s)}{1 + \lambda \hat{\xi}_{i,s}(\beta_s)} = 0. \tag{11}$$

Theorem 2. *Suppose that Conditions* C_1–C_5 *hold. If* β_s *is the true value of the parameter, then*

$$\widehat{\mathcal{R}}_{n,s}(\beta_s) \xrightarrow{d} \chi_1^2, \tag{12}$$

where $\xrightarrow{d}$ *stands for convergence in distribution, and* χ_1^2 *is a standard chi-square random variables with* 1 *degree of freedom.*

Let $\chi_1^2(1-\alpha)$ be the $1-\alpha$ quantile of $\chi_1^2 (0 < \alpha < 1)$. By Theorem 2 an approximate $1-\alpha$ confidence interval for β_s can be defined by $\mathcal{I}_\alpha(\tilde{\beta}_s) = \{\tilde{\beta}_s : \widehat{\mathcal{R}}_{n,s}(\tilde{\beta}_s) \leqslant \chi_1^2(1-\alpha)\}$.

Remark 1. The Result of Theorem 2 can be used to testing the single component of β by applying the duality between confidence intervals and hypothesis tests. Some simulation results will be given in Section 3. Relevant discussion can be found in [20].

2.2. Corrected empirical likelihood for $\eta(\cdot)$

For a given β, we define an auxiliary random vector

$$\varphi_i(\eta(x)) = [Y_i - \eta(x)]K((\beta^T X_i - x)/b), \tag{13}$$

where b is a bandwidth. By (2), we have $E\{\varphi_i(\eta(x))\} = 0$. Similar to [18], we can define an empirical log-likelihood ratio function for $\eta(x)$. However, $\varphi_i(\eta(x))$ cannot be directly used as inference for $\eta(x)$ since $\varphi_i(\eta(x))$ contains unknown parameters β. In order to get the empirical likelihood ratio function for $\eta(x)$, we substitute β with $\hat{\beta}$, which is defined in Section 2.1.

Therefore, with the similar argument to [18], one estimator for $\varphi_i(\eta(x))$ can be obtained and an estimated corrected empirical log-likelihood ratio can be defined as

$$\widehat{\mathcal{L}}_n(\eta(x)) = -2\max\left\{\sum_{i=1}^n \log(np_i) : p_i \geqslant 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i\hat{\varphi}_i(\eta(x)) = 0\right\}, \tag{14}$$

where

$$\hat{\varphi}_i(\eta(x)) = \{Y_i - \eta(x) - [\hat{\eta}(\hat{\beta}^T X_i) - \hat{\eta}(x)]\}K((\hat{\beta}^T X_i - x)/b)$$

and $\hat{\eta}(x) = \hat{\eta}(x; \hat{\beta})$.

By the Lagrange multiplier method, $\widehat{\mathcal{L}}_n(\eta(x))$ can be represented as

$$\widehat{\mathcal{L}}_n(\eta(x)) = 2\sum_{i=1}^n \log\{1 + \theta\hat{\varphi}_i(\eta(x))\}, \tag{15}$$

where θ is determined by

$$\frac{1}{n}\sum_{i=1}^n \frac{\hat{\varphi}_i(\eta(x))}{1+\theta\hat{\varphi}_i(\eta(x))} = 0. \tag{16}$$

Theorem 3. *Suppose that Conditions* C_1–C_6 *in the Appendix hold. For given x, if $\eta(x)$ is the true value of the parameter, then*

$$\widehat{\mathcal{L}}_n(\eta(x)) \xrightarrow{d} \chi_1^2. \tag{17}$$

Let $\chi_1^2(1-\alpha)$ be the $1-\alpha$ quantile of $\chi_1^2 (0 < \alpha < 1)$. By Theorem 3 an approximate $1-\alpha$ confidence interval for $\eta(x)$ can be defined by $\mathcal{I}_\alpha(\tilde{\eta}(x)) = \{\tilde{\eta}(x) : \widehat{\mathcal{L}}_n(\tilde{\eta}(x)) \leqslant \chi_1^2(1-\alpha)\}$.

3. Numerical results

3.1. Simulation studies

To illustrate the numerical performance of our proposed method, we conduct a small simulation experiment in which the sample size $n = 50, 100, 150$. We generate the data from the single-index model

$$Y_i = \eta(\beta^T X_i) + \varepsilon_i, \quad i = 1, \ldots, n, \tag{18}$$

where $\beta = (-1, 2, 1, 1, -3)/4$, X_i is a five-dimensional column vector with the independent uniform $(-0.5, 0.5)$ components, and $\varepsilon_i \sim N(0, 0.4^2)$. In the above model, $\eta(x) = x^3$.

In nonparametric regressions, we use the kernel function $K(u) = 0.75(1-u^2)_+$ for all smoothing steps. The estimated bandwidth $\hat{h}_{opt}$ for h can be obtained by the "leave-one-sample-out" method, see [7] and [5]. To obtain the root-n estimators of the parameters, the final estimated bandwidth for h is taken as $n^{-2/15}\hat{h}_{opt}$, because this guarantees that the required bandwidth has correct order of magnitude for the optimal asymptotic performance, see [1].

To evaluate the performance of the proposed methods for the single parameter of β, two methods are compared: the profile empirical likelihood (PEL) method and the normal approximation (NA) method. The average lengths and coverage probabilities of the confidence interval, with the nominal level $1-\alpha = 0.95$ are computed with 1000 simulation runs and all the simulation results are reported in Tables 1 and 2.

Table 1
Average lengths of confidence intervals for $\beta_s, s = 1, 2, 3, 4, 5$ when nominal confidence level 95%.

Methods	$n\backslash$Parameters	β_1	β_2	β_3	β_4	β_5
PEL	50	0.086	0.081	0.090	0.079	0.084
	100	0.071	0.070	0.062	0.076	0.067
	150	0.048	0.035	0.034	0.042	0.032
NA	50	0.112	0.107	0.098	0.105	0.104
	100	0.079	0.078	0.072	0.089	0.075
	150	0.052	0.042	0.041	0.050	0.049

Table 2
Coverage probabilities of confidence intervals for $\beta_s, s = 1, 2, 3, 4, 5$ when nominal confidence level 95%.

Methods	$n\backslash$Parameters	β_1	β_2	β_3	β_4	β_5
PEL	50	0.912	0.914	0.910	0.918	0.916
	100	0.927	0.923	0.931	0.925	0.921
	150	0.938	0.934	0.940	0.936	0.941
NA	50	0.904	0.898	0.901	0.918	0.902
	100	0.917	0.910	0.928	0.923	0.918
	150	0.932	0.928	0.939	0.930	0.942

Table 3
Average lengths of confidence intervals for link function $\eta(x)$ at the five selected points when nominal confidence level 95%.

Methods	$n \backslash x$	-0.4	-0.3	0.1	0.3	0.4
PEL	50	0.441	0.453	0.441	0.432	0.456
	100	0.315	0.342	0.332	0.315	0.364
	150	0.185	0.177	0.167	0.165	0.193
NA	50	0.485	0.475	0.458	0.449	0.463
	100	0.334	0.357	0.347	0.326	0.370
	150	0.198	0.204	0.194	0.171	0.198

Table 4
Coverage probabilities of confidence intervals for link function $\eta(x)$ at the five selected points when nominal confidence level 95%.

Methods	$n \backslash x$	-0.4	-0.3	0.1	0.3	0.4
PEL	50	0.901	0.910	0.902	0.899	0.903
	100	0.920	0.919	0.923	0.921	0.927
	150	0.931	0.940	0.941	0.936	0.932
NA	50	0.884	0.895	0.886	0.890	0.897
	100	0.917	0.910	0.920	0.918	0.928
	150	0.928	0.941	0.936	0.928	0.927

From the above simulation results, we draw the following conclusions. From Tables 1 and 2, it is easy to see that the PEL performs much better than the NA in terms of coverage probabilities and the average lengths of the confidence intervals. The coverage probabilities of the PEL and NA methods are close to the nominal level as n increases, and from Table 2, we see that the coverage probabilities increase as the n increases.

We also show the performance of the corrected empirical likelihood for the link function $\eta(\cdot)$ in terms of the average lengths and coverage probabilities of the pointwise confidence intervals in Tables 1 and 3. Here the estimated bandwidth for h is taken as $n^{-2/15}\hat{h}_{opt}$, and to save computational time, we tried the simple estimated bandwidth $\hat{b} = an^{-1/5}$ for $a = 0.25, 0.75, 1, 1.5, 2$, which satisfy the condition in Theorem 3. The final bandwidth $\hat{b} = 1.5n^{-1/5}$. The numerical results are fairly stable against shifting values of the selected bandwidth. The kernel function $K(u) = 0.75(1 - u^2)_+$ was also used. We take the sample size as $n = 50, 100, 150$, respectively. In each case the number of simulated realizations is 1000. The pointwise confidence intervals for the nonparametric function $\eta(\cdot)$ at the selected four points $x = -0.4, -0.3, 0.1, 0.3$ and 0.4 are presented in Tables 3 and 4. The average lengths of the confidence intervals are given in Table 3, and the corresponding coverage probabilities are also presented in Table 4.

Table 3 shows that the average interval lengths decrease as the sample size increases and the lengths of the PEL-based intervals are slightly shorter than those based on the NA method. From Table 4 we may conclude that the coverage probabilities based on the PEL method are mostly closer to the nominal level than those based on the NA method.

In addition, based on the PEL method we also considered a simple calculation on the power of the tests in the above example. Here we take the null hypothesis $H_0 : \beta_1 = 0$ and the power functions are evaluated under a sequence of the

Table 5
Powers of the tests based on the PEL method.

$n\backslash\delta$	0	0.010	0.015	0.020	0.025	0.030	0.035
50	0.013	0.124	0.305	0.689	0.836	0.997	1.000
100	0.009	0.148	0.325	0.699	0.890	0.998	1.000
150	0.011	0.208	0.355	0.710	0.954	1.000	1.000

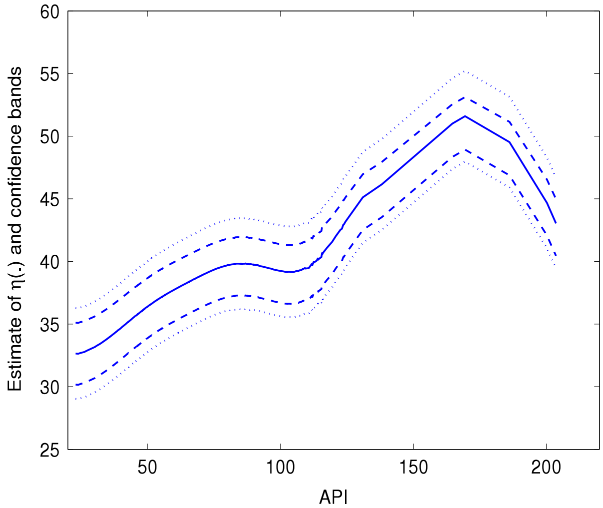

Fig. 1. Application to environmental data. The estimated curves of $\eta(\cdot)$ and its confidence bands based on the proposed empirical likelihood (dashed curve) and the NA method (dotted curve).

alternative models indexed by $\delta : H_1 : \beta_1 = n^{-1/2}\delta,\ \delta = 0, 0.010, 0.015, 0.020, \ldots$. The other case can be investigated in a similar fashion. More discussion can be found in [3] and [6]. The simulation results are reported in Table 5, which depicts the power function based on 1000 simulations at the significance levels: $\gamma = 0.01$, and $n = 50, 100, 150$. When $\delta = 0$, the special alternative collapses into the null hypothesis. Based on the PEL method and different sample sizes, the powers at $\delta = 0$ for the foregoing significance level are 0.013, 0.009, 0.011 respectively. This shows that the PEL gives the right levels of tests under different sample sizes. And the three power functions increases rapidly as δ increases. This shows that the proposed PEL works well.

3.2. A real example

We illustrate the proposed method by an application to the environmental data set, which consists of daily average measurements of pollutants in Hong Kong between January 1, 2000 and December 31, 2000. The description of the data set is given in the introduction. We in this paper employed the single-index model to fit the given data. The single-index model can be expressed as

$$Y = \eta(\beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4) + \varepsilon. \tag{19}$$

On the basis of the empirical likelihood ratio test procedure given in Remark 1, the hypothesis "$\beta_l = 0$" for every $l = 1, 2, 3, 4$ is not significant at the significance level 0.01. So we employ the single-index model (19) to fit the given data set. The kernel function and the bandwidth selection method given in simulation example were used. And the estimate of $(\beta_1, \beta_2, \beta_3, \beta_4)$ is $(0.210, 0.156, 0.920, 0.294)$. For every $l = 1, 2, 3, 4$, the 95% confidence intervals for β_l based on the normal approximation are (−0.181,0.601), (−0.157,0.469), (0.458,1.382) and (−0.093,0.681), respectively. And their 95 confidence intervals based on the proposed empirical likelihood methods are (0.086,0.582), (0.041,0.379), (0.549,1.108), and (0.032,0.547), respectively. These results indicate that, for this data set, the PEL-based confidence intervals is shorter than that based on the NA. The estimated link function and its confidence intervals based on the PEL and the NA are presented in Fig. 1. And Fig. 1 also shows that the confidence band based on the proposed empirical likelihood method is narrower than that based on the NA method.

Acknowledgments

The author thanks the Editor, the Associate Editor, and two referees for their constructive comments, which substantially improved an earlier version of this paper. This research was supported in part by the National Natural Science Foundation of China (10871072), Ph.D. Program Foundation of Ministry of Education of China (20090076110001).

Appendix

Before we prove the main theorems, we now give a set of conditions used in this paper.

Conditions

(C_1) $\eta(\cdot)$ and $\mu(\cdot)$ have Lipschitz continuous second derivative, and the density of $\beta^T X, f(t)$, is bounded away from zero and satisfies the Lipschitz condition of order 1 on $\mathcal{T} = \{t = \beta^T x : x \in \mathcal{A}\}$, and $\mathcal{A}$ is a compact support of X.

(C_2) The kernel $K(t)$ is a bounded and symmetric probability density function and is twice continuously differentiable at t, and satisfies $\int t^2 K(t)\mathrm{d}t \neq 0$ and $\int |t|^k K(t)\mathrm{d}t < \infty,\ k = 1, 2, \ldots$.

(C_3) $\sup_x\{E(\varepsilon^2|\mathbf{X})\} < \infty, \sup_x\{E(\varepsilon^4|\mathbf{X} = x)\} < \infty$.

(C_4) (i)$V = E\{\eta'(\beta^T X)^2 \mathbf{J}^T_{\beta^{(r)}}(X - E(X|\beta^T X))(X - E(X|\beta^T X))^T \mathbf{J}_{\beta^{(r)}}\}$ is a positive definite matrix, where $\mathbf{J}_{\beta^{(r)}}$ is defined in Section 2.

(C_5) The bandwidth h satisfies that $nh^4 \to 0$ and $nh^3 \to \infty$ as $n \to \infty$.

(C_6) The bandwidth b satisfies that $b = cn^{-1/5}$, where c is some positive constant.

In this section, to simplify the notations, we write $\hat{\eta}(\hat{\beta}^T_{/s}X_i) = \hat{\eta}(\hat{\beta}^T_{/s}X_i;\, \hat{\beta}_{/s})$, $\hat{\eta}'(\hat{\beta}^T_{/s}X_i) = \hat{\eta}'(\hat{\beta}^T_{/s}X_i;\, \hat{\beta}_{/s})$ and $\hat{\mu}(\hat{\beta}^T_{/s}X_i) = \hat{\mu}(\hat{\beta}^T_{/s}X_i;\, \hat{\beta}_{/s})$. We firstly present the proof of Theorem 1.

Proof of Theorem 1. By using the similar method to that for Theorem 1 of [2], we can complete the proof of Theorem 1. Here, to save the space, we omit its details. □

To prove Theorem 2, it suffices to prove the following Lemmas.

Lemma 1. *Under the assumptions of Theorem 2, if $\eta(x)$ is the true value of the parameter, we have*

$$\text{(i)}\ \frac{1}{\sqrt{n}}\sum_{i=1}^n \hat{\xi}_{i,s}(\beta_s) \xrightarrow{d} N(0, \Sigma_s), \qquad \text{(ii)}\ \frac{1}{n}\sum_{i=1}^n \hat{\xi}_{i,s}(\beta_s)\hat{\xi}_{i,s}(\beta_s)^T \xrightarrow{p} \Sigma_s,$$

where $\Sigma_s = E\{\eta'(\beta^T X)^2 e_s^T \mathbf{J}^T_{\beta^{(r)}}(X - E(X|\beta^T X))(X - E(X|\beta^T X))^T \mathbf{J}_{\beta^{(r)}} e_s\}$.

Proof. (i) By the definition of $\hat{\xi}_{i,s}(\beta_s)$, we can obtain that

$$\begin{aligned}
\hat{\xi}_{i,s}(\beta_s) &= \eta'(\beta^T X_i)e_s^T \mathbf{J}^T_{\beta^{(r)}}(X_i - \mu(\beta^T X_i))\varepsilon_i + \eta'(\beta^T X_i)e_s^T(\mathbf{J}_{\hat{\beta}^{(r)}_{/s}} - \mathbf{J}_{\beta^{(r)}})^T(X_i - \hat{\mu}(\hat{\beta}^T_{/s}X_i))\varepsilon_i \\
&\quad + \eta'(\beta^T X_i)e_s^T \mathbf{J}^T_{\beta^{(r)}}(\mu(\beta^T X_i) - \hat{\mu}(\hat{\beta}^T_{/s}X_i))\varepsilon_i + (\eta(\beta^T X_i) - \hat{\eta}(\hat{\beta}^T_{/s}X_i))\eta'(\beta^T X_i)e_s^T \mathbf{J}^T_{\hat{\beta}^{(r)}_{/s}}(X_i - \hat{\mu}(\hat{\beta}^T_{/s}X_i)) \\
&\quad + (\eta(\beta^T X_i) - \hat{\eta}(\hat{\beta}^T_{/s}X_i))(\hat{\eta}'(\hat{\beta}^T_{/s}X_i) - \eta'(\beta^T X_i))e_s^T \mathbf{J}^T_{\hat{\beta}^{(r)}_{/s}}(X_i - \hat{\mu}(\hat{\beta}^T_{/s}X_i)) \\
&\quad + (\hat{\eta}'(\hat{\beta}^T_{/s}X_i) - \eta'(\beta^T X_i))e_s^T \mathbf{J}^T_{\hat{\beta}^{(r)}_{/s}}(X_i - \hat{\mu}(\hat{\beta}^T_{/s}X_i))\varepsilon_i \\
&\equiv \Theta_{i1} + \Theta_{i2} + \Theta_{i3} + \Theta_{i4} + \Theta_{i5} + \Theta_{i6}.
\end{aligned} \tag{A.1}$$

By the central limit theorem with condition (C_4), we have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n \Theta_{i1} \xrightarrow{d} N(0, \Sigma_s). \tag{A.2}$$

Therefore, to prove Lemma 1(i), we only need to show that $\frac{1}{\sqrt{n}}\sum_{i=1}^n \Theta_{il} \xrightarrow{p} 0,\ l = 2, 3, 4, 5, 6$.

By a Taylor expansion, we have

$$\hat{\mu}(\hat{\beta}^T_{/s}X_i) - \mu(\beta^T X_i) = \mu'(\beta^T X_i)(\hat{\beta}_{/s} - \beta)^T X_i + \hat{\mu}(\beta^T X_i) - \mu(\beta^T X_i) + o_p(n^{-1/2}), \tag{A.3}$$

and from Theorem 1 we can obtain that

$$\mathbf{J}_{\hat{\beta}^{(r)}_{/s}} - \mathbf{J}_{\beta^{(r)}} \xrightarrow{p} 0, \qquad \hat{\beta}_{/s} - \beta = O_p(n^{-1/2}). \tag{A.4}$$

Note that

$$\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{i=1}^n \Theta_{i2} &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \eta'(\beta^T X_i)e_s^T(\mathbf{J}_{\hat{\beta}^{(r)}_{/s}} - \mathbf{J}_{\beta^{(r)}})^T(X_i - \mu(\beta^T X_i))\varepsilon_i \\
&\quad + \frac{1}{\sqrt{n}}\sum_{i=1}^n \eta'(\beta^T X_i)e_s^T(\mathbf{J}_{\hat{\beta}^{(r)}_{/s}} - \mathbf{J}_{\beta^{(r)}})^T \mu'(\beta^T X_i)(\beta - \hat{\beta}_{/s})^T X_i\varepsilon_i
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta'(\beta^TX_i)e_s^T(\mathbf{J}_{\hat\beta_{/s}^{(r)}}-\mathbf{J}_{\beta^{(r)}})^T(\mu(\beta^TX_i)-\hat\mu(\beta^TX_i))\varepsilon_i\\
&+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta'(\beta^TX_i)e_s^T(\mathbf{J}_{\beta^{(r)}}-\mathbf{J}_{\hat\beta_{/s}^{(r)}})^To_p(n^{-1/2})\varepsilon_i\\
\equiv{}& A_1+A_2+A_3+A_4.
\end{aligned}
$$

By using the similar arguments to those used in Lemma 4 of [21] and the (A.4), we can easily prove that $A_k \xrightarrow{p} 0,\ k = 1, 2, 3, 4$. This implies that $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Theta_{i2} \xrightarrow{p} 0$. Similarly, we have $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Theta_{i3} \xrightarrow{p} 0$.

Next we consider $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Theta_{i4}$.

Similar to (A.3), we have

$$
\hat\eta(\hat\beta_{/s}^TX_i)-\eta(\beta^TX_i)=\eta'(\beta^TX_i)(\hat\beta_{/s}-\beta)^TX_i+\hat\eta(\beta^TX_i)-\eta(\beta^TX_i)+o_p(n^{-1/2}). \tag{A.5}
$$

Thus we have

$$
\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Theta_{i4}={}&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta'(\beta^TX_i)^2(\beta-\hat\beta_{/s})^TX_ie_s^T\mathbf{J}^T_{\hat\beta_{/s}^{(r)}}(X_i-\hat\mu(\hat\beta_{/s}^TX_i))\\
&+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta'(\beta^TX_i)(\eta(\beta^TX_i)-\hat\eta(\beta^TX_i))e_s^T\mathbf{J}^T_{\hat\beta_{/s}^{(r)}}(X_i-\hat\mu(\hat\beta_{/s}^TX_i))\\
&+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta'(\beta^TX_i)o_p(n^{-1/2})e_s^T\mathbf{J}^T_{\hat\beta_{/s}^{(r)}}(X_i-\hat\mu(\hat\beta_{/s}^TX_i))\\
\equiv{}& B_1+B_2+B_3.
\end{aligned} \tag{A.6}
$$

Similar to $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Theta_{i2}$, we have $B_j \xrightarrow{p} 0,\ j = 1, 3$. As to B_2, we have

$$
\begin{aligned}
B_2={}&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta'(\beta^TX_i)(\eta(\beta^TX_i)-\hat\eta(\beta^TX_i))e_s^T\mathbf{J}^T_{\beta^{(r)}}(X_i-\hat\mu(\beta^TX_i))\\
&+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta'(\beta^TX_i)(\eta(\beta^TX_i)-\hat\eta(\beta^TX_i))e_s^T\mathbf{J}^T_{\beta^{(r)}}[(\mu'(\beta^TX_i)(\beta-\hat\beta_{/s})^TX_i)+o_p(n^{-1/2})]\\
&+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta'(\beta^TX_i)(\eta(\beta^TX_i)-\hat\eta(\beta^TX_i))e_s^T(\mathbf{J}_{\hat\beta_{/s}^{(r)}}-\mathbf{J}_{\beta^{(r)}})^T(X_i-\hat\mu(\beta^TX_i))\\
&+\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta'(\beta^TX_i)(\eta(\beta^TX_i)-\hat\eta(\beta^TX_i))e_s^T(\mathbf{J}_{\hat\beta_{/s}^{(r)}}-\mathbf{J}_{\beta^{(r)}})^T[(\mu'(\beta^TX_i)(\beta-\hat\beta_{/s})^TX_i)+o_p(n^{-1/2})]\\
\equiv{}& B_{21}+B_{22}+B_{23}+B_{24}.
\end{aligned} \tag{A.7}
$$

By using similar arguments to that for Lemma 4 and the (A.4), we can easily show that $B_{2l} \xrightarrow{p} 0,\ l = 1, 2, 3, 4$, which implies that $B_2 \xrightarrow{p} 0$. Therefore, we have $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Theta_{i4} \xrightarrow{p} 0$. Similarly, we can also show that $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Theta_{i5} \xrightarrow{p} 0$ and $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Theta_{i6} \xrightarrow{p} 0$. Therefore we complete the proof of Lemma 1(i).

(ii) We also use the notations in the proof of this term. Let $\Theta_i^\star = \Theta_{i2} + \Theta_{i3} + \Theta_{i4} + \Theta_{i5} + \Theta_{i6}$. Thus we have

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\hat\xi_{i,s}(\beta_s)\hat\xi_{i,s}(\beta_s)^T={}&\frac{1}{n}\sum_{i=1}^{n}\Theta_{i1}\Theta_{i1}^T+\frac{1}{n}\sum_{i=1}^{n}\Theta_{i1}\Theta_i^{\star T}+\frac{1}{n}\sum_{i=1}^{n}\Theta_i^\star\Theta_{i1}^T+\frac{1}{n}\sum_{i=1}^{n}\Theta_i^\star\Theta_i^{\star T}\\
\equiv{}&\Upsilon_1+\Upsilon_2+\Upsilon_3+\Upsilon_4.
\end{aligned}
$$

By the law of large numbers, we can derive that $\Upsilon_1 \xrightarrow{p} \Sigma_s$. By using the similar argument to that for Lemma 5 of [21], we can show that $\Upsilon_k \xrightarrow{p} 0,\ k = 2, 3, 4$. This completes the proof. □

Lemma 2. *Under the assumptions of Theorem 1, if $\eta(x)$ is the true value of the parameter, we have*

$$
\text{(i)}\ \max_{1\leqslant i\leqslant n}|\hat\xi_{i,s}(\beta_s)|=o_p(n^{1/2}),\qquad \text{(ii)}\ |\lambda|=O_p(n^{-1/2}).
$$

Proof. (i) By the fact that for any sequence of independent and identically distributed random variables $\zeta_i,\ i = 1, \ldots, n$ with $E(\zeta_i^T\zeta_i) < \infty$, we have $\max_{1\leqslant i\leqslant n}\|\zeta_i\| = o_p(n^{1/2})$, which leads to $\max_{1\leqslant i\leqslant n}|\Theta_{i1}| = o_p(n^{1/2})$. Applying the similar techniques as were used in the analysis of Lemma 6 of [21], we have $\max_{1\leqslant i\leqslant n}|\Theta_{ik}| = o_p(n^{1/2}),\ k = 2, 3, 4, 5, 6$. This entails the result.

(ii) Together with Lemma 1(ii) and using the arguments similar to [11], we can obtain $|\lambda| = O_p(n^{-1/2})$. □

Proof of Theorem 2. Applying a Taylor series expansion to Eq. (10) and invoking Lemmas 1 and 2, we can obtain that

$$\widehat{\mathcal{R}}_{n,s}(\beta_s) = 2\sum_{i=1}^{n}\left[\lambda\hat{\xi}_{i,s}(\beta_s) - \frac{1}{2}(\lambda\hat{\xi}_{i,s}(\beta_s))^2\right] + o_p(1). \tag{A.8}$$

By Eq. (8), it follows that

$$\begin{aligned}0 &= \frac{1}{n}\sum_{i=1}^{n}\frac{\hat{\xi}_{i,s}(\beta_s)}{1+\lambda\hat{\xi}_{i,s}(\beta_s)}\\ &= \frac{1}{n}\sum_{i=1}^{n}\hat{\xi}_{i,s}(\beta_s) - \frac{1}{n}\sum_{i=1}^{n}\hat{\xi}_{i,s}(\beta_s)[\hat{\xi}_{i,s}(\beta_s)]^T\lambda + \frac{1}{n}\sum_{i=1}^{n}\frac{\hat{\xi}_{i,s}(\beta_s)[\lambda\hat{\xi}_{i,s}(\beta_s)]^2}{1+\lambda\hat{\xi}_{i,s}(\beta_s)}.\end{aligned}$$

By using Lemmas 2–4, we obtain

$$\sum_{i=1}^{n}[\lambda\hat{\xi}_{i,s}(\beta_s)]^2 = \sum_{i=1}^{n}\lambda\hat{\xi}_{i,s}(\beta_s) + o_p(1), \tag{A.9}$$

$$\lambda = \left[\sum_{i=1}^{n}\hat{\xi}_{i,s}(\beta_s)(\hat{\xi}_{i,s}(\beta_s))^T\right]^{-1}\sum_{i=1}^{n}\hat{\xi}_{i,s}(\beta_s) + o_p(n^{-1/2}). \tag{A.10}$$

Then, by Eqs. (A.3)–(A.5), we have

$$\widehat{\mathcal{R}}_{n,s}(\beta_s) = \left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\hat{\xi}_{i,s}(\beta_s)\right]^T\left[\frac{1}{n}\sum_{i=1}^{n}\hat{\xi}_{i,s}(\beta_s)(\hat{\xi}_{i,s}(\beta_s))^T\right]^{-1}\left[\sum_{i=1}^{n}\hat{\xi}_{i,s}(\beta_s)\right] + o_p(1).$$

This, together with Lemmas 1 and 2, completes the proof of Theorem 2. □

Lemma 3. *Under the assumptions of Theorem 3, if $\eta(x)$ is the true value of the parameter, we have*

$$\text{(i)}\ \frac{1}{\sqrt{nb}}\sum_{i=1}^{n}\hat{\varphi}_i(\eta(x)) \xrightarrow{d} N(0, \Sigma(x)), \qquad \text{(ii)}\ \frac{1}{nb}\sum_{i=1}^{n}\hat{\varphi}_i(\eta(x))\hat{\varphi}_i^T(\eta(x)) \xrightarrow{p} \Sigma(x),$$

where $\Sigma(x) = \sigma^2(x)f(x)\int K^2(t)\mathrm{d}t$, $\sigma^2(x) = E(\varepsilon^2|x)$.

Proof. By Part (i) and using the similar arguments as those used in Lemma 5 of [21], one can complete the proof of Part (ii). Here we only prove (i).

By the definition of $\hat{\varphi}_i(\eta(x))$ and the Taylor expansion, it is easy to show that

$$\begin{aligned}\frac{1}{\sqrt{nb}}\sum_{i=1}^{n}\hat{\varphi}_i(\eta(x)) &= \frac{1}{\sqrt{nh}}\sum_{i=1}^{n}\{Y_i - \eta(\beta^TX_i)\}K((\hat{\beta}^TX_i - x)/b)\\ &\quad + \frac{1}{\sqrt{nh}}\sum_{i=1}^{n}\{(\eta(\beta^TX_i) - \hat{\eta}(\hat{\beta}^TX_i) + (\hat{\eta}(x) - \eta(x)))\}K((\hat{\beta}^TX_i - x)/b) + o_p(1)\\ &\equiv \Delta_1 + \Delta_2 + o_p(1).\end{aligned} \tag{A.11}$$

Noting that, for every i

$$K((\hat{\beta}^TX_i - x)/b) = K((\beta^TX_i - x)/b) + K'((\beta^TX_i - x)/b)(\hat{\beta} - \beta)^TX_i + o_p(n^{-1/2}), \tag{A.12}$$

which, combining with the result that $\hat{\beta} - \beta = O_p(n^{-1/2})$ from Theorem 1 and simple calculation, leads to

$$\Delta_2 = \frac{1}{\sqrt{nh}}\sum_{i=1}^{n}\{Y_i - \eta(\beta^TX_i)\}K((\beta^TX_i - x)/b) + o_p(1). \tag{A.13}$$

Using the central limits theorem, we have

$$\Delta_2 \xrightarrow{d} N(0, \Sigma(x)). \tag{A.14}$$

Next we prove

$$\Delta_2 \xrightarrow{p} 0. \tag{A.15}$$

Noting that

$$\hat{\eta}(\hat{\beta}^T X_i) - \eta(\beta^T X_i) = \eta'(\beta^T X_i)(\hat{\beta} - \beta)^T X_i + \hat{\eta}(\beta^T X_i) - \eta(\beta^T X_i) + o_p(n^{-1/2}).$$

By this, it can be shown that

$$\begin{aligned}
\Delta_2 &= \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \Big\{(\hat{\eta}(x) - \eta(x)) - \eta'(\beta^T X_i)(\hat{\beta} - \beta)^T X_i - (\hat{\eta}(\beta^T X_i) - \eta(\beta^T X_i)) + o_p(n^{-1/2})\Big\} K((\hat{\beta}^T X_i - x)/b) \\
&= \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \{\hat{\eta}(x) - \eta(x)\} K((\hat{\beta}^T X_i - x)/b) + \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \{\hat{\eta}(\beta^T X_i) - \eta(\beta^T X_i)\} K((\hat{\beta}^T X_i - x)/b) \\
&\quad + \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \{\eta'(\beta^T X_i)(\hat{\beta} - \beta)^T X_i\} K((\hat{\beta}^T X_i - x)/b) + \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} K((\hat{\beta}^T X_i - x)/b) o_p(n^{-1/2}) \\
&\equiv \Delta_{31} + \Delta_{32} + \Delta_{33} + \Delta_{34}.
\end{aligned} \tag{A.16}$$

By using (A.12) and the similar arguments to those used in Lemmas 3 and 4 of [21], we can obtain that

$$\Delta_{3k} \xrightarrow{p} 0, \quad k = 1, 2, 3, 4. \tag{A.17}$$

Thus we prove (A.15). Therefore, we complete the proof of Lemma 3. □

Lemma 4. *Under the assumptions of Theorem 1, if $\eta(x)$ is the true value of the parameter, we have*

$$\text{(i)} \max_{1 \leqslant i \leqslant n} \|\hat{\varphi}_i(\eta(x))\| = o_p((nb)^{1/2}), \qquad \text{(ii)}\ \theta = O_p((nb)^{-1/2}).$$

Proof. As to part (ii), by Lemma 3(i), it can be shown that $\frac{1}{nb} \sum_{i=1}^{n} \hat{\varphi}_i(\eta(x)) = O_p((nb)^{-1/2})$, which, combining with Lemma 3(ii), leads to Lemma 4(ii) by using the same arguments as that used in the proof of expression (2.14) in [11]. Next we prove part (i).

Let $\tilde{\Lambda}_i = \varepsilon_i K((\hat{\beta}^T X_i - x)/b)$ and denote $\tilde{\eta}_i = (\hat{\eta}(\hat{\beta}^T X_i) - \eta(\beta^T X_i)) K((\hat{\beta}^T X_i - x)/b)$, $\tilde{\eta}_{ix} = (\hat{\eta}(x) - \eta(x)) K((\hat{\beta}^T X_i - x)/b)$. We also use the notation of Lemma 3 and it is easy to show that

$$\begin{aligned}
\max_{1 \leqslant i \leqslant n} |\hat{\varphi}_i(\eta(x))| &\leqslant c \max_{1 \leqslant i \leqslant n} |\tilde{\Lambda}_i| + c \max_{1 \leqslant i \leqslant n} |\tilde{\eta}_i| + c \max_{1 \leqslant i \leqslant n} |\tilde{\eta}_{ix}| \\
&\equiv \Xi_1 + \Xi_2 + \Xi_3,
\end{aligned}$$

where $c > 0$ generally represents any constant which may take a different value for each appearance. To prove Lemma 4(i), we only need to prove that $\Xi_k = o_p(n^{1/2})$, $k = 1, 2, 3$.

Similarly to the proof of the above Lemma 3 and the proof of Lemma A.1 in [18], we can obtain that $\Xi_l = o_p((nb)^{1/2})$, $l = 1, 2, 3$. Thus, the proof of Lemma 4 is completed. □

Proof of Theorem 3. Applying a Taylor series expansion to Eq. (15) and invoking Lemmas 3 and 4, we can obtain that

$$\widehat{\mathcal{L}}_n(\eta(x)) = \left(\frac{1}{\sqrt{nb}} \sum_{i=1}^{n} \hat{\varphi}_i(\eta(x))\right)^T \left(\frac{1}{nb} \sum_{i=1}^{n} \hat{\varphi}_i(\eta(x)) \hat{\varphi}_i^T(\eta(x))\right)^{-1} \left(\frac{1}{\sqrt{nb}} \sum_{i=1}^{n} \hat{\varphi}_i(\eta(x))\right) + o_p(1).$$

With the similar argument to the proof of Theorem 2, we complete the proof of Theorem 3. □

References

[1] R.J. Carroll, J. Fan, I. Gijbels, M.P. Wand, Generalized partially linear single-index models, J. Amer. Statist. Assoc. 92 (1997) 477–489.
[2] Z. Chang, L. Xue, L. Zhu, On an asymptotically more efficient estimation of the single-index model, J. Multivariate Anal. 101 (2010) 1898–1901.
[3] S.X. Chen, Comparing empirical likelihood and bootstrap hypothesis tests, J. Multivariate Anal. 51 (1994) 277–293.
[4] X. Chen, H.J. Cui, Empirical likelihood for partially linear single-index errors-in-variables model, Comm. Statist. Theory Methods 38 (2009) 2498–2514.
[5] J. Fan, I. Gijbels, Local Polynomial Modeling and its Applications, Chapman and Hall, London, 1996.
[6] P. Hall, B. La Scala, Methodology and algorithms of empirical likelihood, Internat. Statist. Rev. 58 (1990) 109–127.
[7] W. Härdle, P. Hall, H. Ichimura, Optimal smoothing in single-index models, Ann. Statist. 21 (1993) 157–178.
[8] H. Ichimura, Semiparametric least squares (SLS) and weighted SLS estimation of singleindex models, J. Econometrics 58 (1993) 71–120.
[9] H. Liang, N.S. Wang, Partially linear single-index measurement error models, Statist. Sinica 15 (2005) 99–116.
[10] A.B. Owen, Empirical likelihood ratio confidence intervals for a single function, Biometrika 75 (1988) 237–249.
[11] A.B. Owen, Empirical likelihood ratio confidence regions, Ann. Statist. 18 (1990) 90–120.
[12] W. Stute, L.X. Zhu, Nonparametric checks for single-index models, Ann. Statist. 33 (2005) 1048–1083.

[13] J.L. Wang, L.G. Xue, L.X. Zhu, Y.S. Chong, Estimation for a partial-linear single-index model, Ann. Statist. 38 (2010) 246–274.
[14] Y. Xia, W. Härdle, Semi-parametric estimation of partially linear single-index models, J. Multivariate Anal. 97 (2006) 1162–1184.
[15] Y. Xia, H. Tong, W.K. Li, On extended partially linear single-index models, Biometrika 86 (1999) 831–842.
[16] Y. Xia, H. Tong, W.K. Li, L.X. Zhu, An adaptive estimation of dimension reduction space, J. Roy. Statist. Assoc. B 64 (2002) 363–410.
[17] L.G. Xue, L.X. Zhu, Empirical likelihood for single-index model, J. Multivariate Anal. 97 (2006) 1295–1312.
[18] L.G. Xue, L.X. Zhu, Empirical likelihood for a varying coefficient model with longitudinal data, J. Amer. Statist. Assoc. 102 (2007) 642–654.
[19] Y. Yu, D. Ruppert, Penalized spline estimation for partially linear single-index models, J. Amer. Statist. Assoc. 97 (2002) 1042–1054.
[20] Z.G. Zhou, W.M. Qian, C. He, Testing serial correlation for partially nonlinear models, J. Korean Statist. Soc. 39 (2010) 501–509.
[21] L.X. Zhu, L.G. Xue, Empirical likelihood confidence regions in a partially linear single-index model, J. Roy. Statist. Assoc. B 68 (2006) 549–570.